From: [redacted]
Sent: Saturday, June 23, 2012 10:19 PM
To: Wessel, Ann (ECY)
Cc: Marguerite Glover
Subject: Formal comment on the Proposed Dungeness Water Management Rule

Version:1.0 StartHTML:0000000168 EndHTML:0000036744 StartFragment:0000000471 EndFragment:0000036727

Dr. Robert N. Crittenden

[redacted]

Ann Wessel
awes461@ecy.wa.gov
1440 10th st., Suite 102
Bellingham WA 98225
360 715-5215

June 23, 2012

Regarding: Formal Comment on the Proposed Dungeness Water Management Rule

Dear Ann,

Please consider these as formal comments. Thank you.

1. The steady-state calibration runs in the 2008 Dungeness Groundwater Flow model have more parameters than data and, therefore, have zero degrees of freedom.

In this comment, I go through the 2008 report showing, where the data and parameters are found and demonstrating that there are more parameters than data. Then, I compute the model's degrees of freedom.

The 2008 Dungeness Groundwater Flow Model, Design, Construction, Calibration, and Results by Clallam County and Pacific Groundwater Group is available on Clallam County's website at clallam.net/environment/assets/applets/PGG_2008_Dungeness_Model_Final_Report.pdf

That is a pdf file. I will refer to the page numbers of the original report as, for example, "Section 5.5.3 on page 35." However, not all the pages of that report were numbered. In particular, the tables and figures at the end of it don't have page numbers. Therefore, for those pages, I will *refer to their location by the page number of the pdf file, for example, "Table 4.2 on pdf page 61."*

They presented the data set that they used in the steady-state calibration runs in their Table 5-3 on pdf page 64 of their report. There are 69 observations. That is,

$$n=69$$

where n is the sample size

They presented the estimated values for the hydraulic conductivity parameters in Table 4-2 on pdf page 61. There are 68 of those parameters.

In particular, that includes 27 vertical conductivities, 27 horizontal conductivities, and twelve vertical conductivities for the streambeds, plus two additional values for the streambeds of Siebert and McDonald Creeks. Thus, 27+27+12+2=68.

However, those were not all the estimated parameters. Other parameters that they estimated are discussed in Section 5.2 beginning on page 27 of their report. There they list most of the parameters that they estimated or adjusted. In particular:

- Aquifer horizontal conductivity --- These were included in the 68 parameters discussed above.
- Aquitard vertical conductivity --- This is at least one additional parameter.
- Streambed vertical conductivity --- These were included in the 68 parameters discussed above.
- Dungeness River Elevation --- As discussed in their text, they treated the Dungeness River as if it were elevated above grade by this fixed amount, in order to better fit the movement of water into and out of its streambed. This is one of the more obvious unrealistic aspects of this model. It adds at least one additional parameter.
- Drain Cell Distributions --- This may have be the locations of the "drain cells", where groundwater finally moves from the ground into saltwater or possibly, alternatively, into a stream. There seem to have been several of these drain cells. They contribute at least one more parameter but probably several.
- Constant head cell vertical conductivity --- There seems to have been at least one of these. That contributes at least one more parameter but possibly more.

These parameters contribute at least four additional parameters but probably more. Taking these additional parameters into account, there were at least 68+4 = 72 estimated parameters. That is more parameters than data.

That is what I wish to show, because, the model's "degrees of freedom", df, is equal to the number of data minus the number of estimated parameters. However, zero is the smallest possible value for the degrees of freedom. So, if there are more parameters than data, the degrees of freedom is set to zero. As that is the case here, the model has zero degrees of freedom:

$$df=0$$

where df is the degrees of freedom

Nevertheless, their model had even more fitted parameters than were included in Table 4-2 or in their explicit list in section 5.2. They are discussed elsewhere in the text. In particular, in the text on page 27 of their report, they discuss varying the horizontal conductivity of the lower aquifer and, also, of one or more additional deep layers. That adds at least two more parameters, bringing the total to at least 68+4+2=74.

Furthermore, by far the largest group of additional parameters that they adjusted to improve the fit of the model were structural parameters, rather that the more usual numerical parameters. Specifically, these were the locations of the boundaries between the various zones in layer 1. They provide a map of those zones in figure 4-10 on pdf page 84. In particular, on page 23 of their report they say, "the boundaries of the various sub-regions were sometimes shifted within layer 1 during calibration." As these changes in the locations of the boundaries could be represented as numerical parameters, they, have to be included as estimated parameters. Potentially, there are several hundred such parameters but they seem to have only adjusted a few of them.

In particular, on page 28 of their report they mention changing the boundaries of Grey's Marsh. There are also several differences in the zone boundaries between realization Dung-7e and Dung-7g. These can be seen by using a light table and overlaying the maps in Figures 4-10 and 4-11 on pfd pages 84 and 85. However, these differences don't reveal what all the changes may have been from the zones, that were originally defined by Dr. Thomas et al. in their 1999 report (Hydrogeological assessment of the Sequim-Dungeness Area. Clallam County, Washington. USGS Water Resourcers Investigation Report 99-4048.)

2. The standard errors for the 2008 model and its parameter estimates are infinite.

In this comment, I compute the standard error, using information provided in the report on the 2008 model.

The standard error of the error components of the observations in a linear model can be estimated using the equation,

$$SE(e) = Sqrt(RSS/df)$$

where:

Sqrt is square-root;
SE() is standard error;
e is the random error component in the observations;
df is the degree of freedom; and

RSS is the residual sum of squares.

You can find that equation in any good book on linear models. For example, Sanford Wiesberg's 1980 book *Applied Linear Regression* by John Wiley and Sons.

The only thing we are lacking to do this computation is the residual sum of squares, RSS. Its value can be found near the bottom of Table 5.3 on pdf page 64 of the report. They give its value for two different realizations of the model. For Dung-7e it is 28523 and for Dung-7g it is 32755. However, it doesn't matter what its value is, so long as it isn't zero. The reason is that, the model has zero degrees of freedom and, consequently, the equation for the standard error involves a division by zero. The answer is infinity.

$$SE(e) = \infty$$

What we just estimated, is the standard error of the error components. However, what we really want is the standard error of the estimated parameters. The equation for that can be found, for example, on page 43 of Weisberg's book (ibid.). However, that introduces a lot of terminology that is trivial if you're familiar with it but is unnecessarily confusing if you aren't. Nevertheless, what it tells you, is that the square of the standard error of the estimated parameters is linear in the square of the standard error of the error components. Consequently, if the latter is infinite, the former is, too.

Before considering the consequences of these results, from comments 1 and 2, I will state a qualification to their applicability.

That is that they are based on the assumption that the groundwater model is a linear model. Groundwater models usually are, because, the relationship between pressure and discharge is linear. However, nonlinear elements could have been inserted into this particular model. I don't know whether that was done or not but it might have been done. However, even if it were, I suspect they had a minor effect on the overall response of the model. Nevertheless, I need to note this qualification.

3. The reason that I didn't directly examine the model's code and its documentation, was that on two occasions, when I requested copies of them, I was told that they were proprietary and I could not have copies.

I was only told that it was not proprietary at the Clallam County Board of Commissioners' hearing on the proposed rule. By then, it was too late for me to obtain and study it. --- I strongly object to the Department's employees telling me a direct lie in order, apparently, to prevent my examining the model.

Incidentally, I did a groundwater model as part of my master's thesis, and several others later on. So, I have little doubt that I could understand it and run it, provided that its code was written transparently and it was properly documented.

4. The subsequent studies that rest upon the static calibration study have the same problems.

If you continue reading the groundwater report beyond the static calibration study, you will find several other studies that were conducted using the parameter values that were estimated in the static calibration study. That is, these subsequent studies were done conditional upon the results of that first study.

However, when you do another study conditional upon an earlier study, you need to incorporate the uncertainty in that earlier study into your estimates of the uncertainty of the second study. You might compute that by, using the conditional variance formula or by, using various alternative approaches. However, in this case the answer is clear: As the standard errors of the first study were infinite, they contribute infinite variation to the results of the subsequent studies.

Those subsequent studies include but are not limited to the transient calibration study, the test of whether the southern boundary is impermeable, the test of whether additional lower layers should be considered in the model, and an aquifer recharge study. Quite a few other studies that used those estimated parameters and/or the results of these secondary studies, were done later on. Obviously, all of them must be considered as being highly suspect.

5. No statistical tests based on the 2008 model can be significant.

Parameters estimated from linear models are normally distributed. (The normal distribution is the familiar bell curve.) Their 95% confidence limits are at the mean plus-or-minus two standard errors. So, if the standard errors are infinite, the 95% confidence limits are at plus-or-minus infinity.

However, if a particular parameter can not have negative values, as is the case for hydraulic conductivities, its confidence limits would be from zero to infinity.

The basic statistical test asks whether an estimate lies beyond the 95% confidence limits. In that case, it is "significant." That means that there is less than a 5% chance that it arose from random variation.

However, if the confidence limits are infinitely broad, it is not possible to ever achieve a significant result. That is the case for the 2008 model, as well as all the secondary studies that were done conditional upon it.

6. Estimates and predictions based on the 2008 model have no scientific support.

If, the parameter estimates can never be significant, because, the model has zero degrees of freedom, there is an unacceptably large chance that they arose solely from random variation. In that case, there is no scientific support for the parameter estimates or any predictions that might be made using that model. --- That is the conclusion that has to be reached for the 2008 groundwater model as well as for all the secondary studies that were done conditional upon it.

7. No meaningful measures of dispersion were computed for the above models.

The accepted scientific practice requires that variances, standard errors, confidence limits and/or other measures of dispersion be computed for estimates. However, those things were not done for the 2008 groundwater study.

Instead, they computed the sample statistics for the observed residuals. They presented them in Table 5-3 on page 64 of their report. They, also, compare them to the range of the residuals. They use those statistics as measures of the accuracy of their results. Unfortunately, that sample statistic is a biased estimator for the dispersion of the error components of the model, because, it does not account for the model's degrees of freedom. The standard errors that I computed above, are the unbiased estimators. You can see from the difference between them, how very large that bias is.

8. The 2008 model contains no statistical tests. Conclusions were reached based on "judgment," instead. That is not a scientifically valid method.

The accepted scientific practice is to conduct statistical tests and to draw reasonable conclusions based upon their results. However, no tests were conducted. The conclusions that were reached appear to have been based on the user's judgment. That would, of course, be expected to include the user's biases and preconceptions. That is contrary to the purpose of science, as science seeks to avoid precisely those types of influences. This shows that the 2008 model is not scientific. The same also applies to the secondary studies that were based upon it.

9. The 2008 model doesn't provide unique solutions.

Because their model has more parameters than data, it will not provide unique estimates. The authors recognized that non-uniqueness and repeatedly mentioned it. For example (on page 47) they said:

"The ability to create more than one model realization capable of meeting calibration criteria (referred to as non-uniqueness) is quite common, and accounts for some of the uncertainty inherent in predicting impacts from hydrologic stresses. This inherent uncertainty can not be avoided in any model or predictive approach, largely because subsurface conditions are inherently variable and data are typically insufficient to characterize such variability. While model predictions can still be performed to obtain estimates of impact at a commonly accepted degree of accuracy, uncertainty associated with non-uniqueness cannot be avoided and prevents prediction of "exact" values of hydrologic impact. In some cases, modelers will use stochastic analysis of multiple (ie. many) realizations to characterize the range of uncertainty in model predictions."

That shows that they recognized the cause of the non-uniqueness but it also reveals their naivety of statistical methods and scientific modeling practices.

If they would reduce the number of parameters in their model and/or get more data, they would obtain unique estimates. Nevertheless, those would still be estimates with random error, rather than exact values.

10. The fitting procedure used in the above models wasn't impartial

They computed two realizations of the parameter values, to provide some impression of how non-uniqueness affected the dispersion in their estimates. However, the values of those two realizations largely reflect choices they made, for they manually searched for optima, and decided where to stop, presumably, when the response surface got relatively flat, instead of using an automated or standardized search procedure.

I must admit that the more widely available automated search procedures don't work very well. They tend to be slow and all too often miss optima or won't converge. Although, more reliable procedures can be written, the usual ones often need manual oversight. However, in doing that, care needs to be taken that the outcome of a search is not the user's choice, as occurred for the two realizations of the groundwater model.

Incidentally, notice that those two realizations have different RSS values. So, clearly they are not two instances of non-unique solutions. They appear to have been selected, instead, from the total range of possible realizations.
They also manually adjusted parameter values in certain cells with large residuals such as dry or flooded cells. However, they seem to have referred back to other reports to correct defects in the model structure for those cells. So, what they did probably was not so serious a flaw in methodology that we would have to conclude that any cell's parameters would have been individually readjusted if they didn't like the way the model fitted it.

Their comments on page A5 of their report, regarding those types of corrections reveal that they did not even consistently hand-adjust the hydraulic conductivities to improve the fit of the model.

Overall, their fitting methodology seems to have been ad-hoc and subject to their personal biases.

11. The "peer review" of the 2008 model wasn't effective.

The "peer review" of the groundwater model appears to have been conducted by individuals who, for the most part, were not independent, scientists, nor peers of the scientific community. However, the courts and the State have diluted the definitions of "peer," "scientist", and "science" so far that they are far removed from the academic standards for peer review.

Also, if the individuals who did the review weren't effectively independent, they might not have commented on any flaws they may have found.

The substantive defect, in their review process, was that none of them criticized the model's having zero degrees of freedom. --- Here, I must admit, that neither did I immediately find that flaw. It is so far from the accepted scientific practice, that it was a long time before it occurred to me that they might have actually written a model with more estimated parameters than data. As a scientist and a modeler, I find that shocking. I would find it beyond belief, that such a thing could be an accepted practice in a licensed profession, except that here it is and some licensed groundwater modelers seem to defend it.

12. The flaws in the groundwater model make it susceptible to abuse and it appears to have been intended to be used for the purposes of arbitrary governance.

I was told by a prominent individual in this County, that Steve Tharinger, who at that time was a Clallam County Commissioner and also chaired WRIA18, had boasted that the groundwater model could provide any outcome that was desired. Furthermore, I was told that he said that before the model was completed. That implies that the model's flaws were known and deliberate and that it was intended to provide a mechanism for arbitrary governance.
I have not named the individual who told me this, because, he is a private person. However, Steve Tharinger is an elected official and it is not slanderous for me to say anything about a public person that I believe is true. --- It is true that I was told this and I suspect that what I was told may be true, too, but whether it is or not is more than I could possibly know.

Although, what I was told is hearsay and, therefore, needs supporting testimony from the individuals who were directly involved, it should be evident from my criticisms of the groundwater model that it could be used that way.

I suggest that if that model is used, despite its flaws, there needs to be a public process conducted by an elected authority to prevent its parameters from being arbitrarily adjusted on a case-by-case or group-by-group basis. That might inhibit its being used for arbitrary governance.

An investigation by law enforcement might, also, be appropriate, to determine whether any violations of the law have already occurred and to prevent any related future violations.

13. The adoption of any new model or upgrades to an existing model should be done by an elected body through an open public process, instead of being under the department's authority.

Section WAC 173-518-070 3ai of the proposed rule says, regarding the 2008 model,

"If ecology determines a better method is available in the future, then ecology will apply the new method."

The concern that is expressed above in criticism #12, is that the non-uniqueness of the 2008 model provides the ability to arbitrary adjust its outcomes and this sentence in the proposed rule, authorizes the department to do precisely that.

It should be replaced by requiring that any change be adopted through an open public process conduced by an elected authority, such as the Clallam County Board of Commissioners. To leave it under the department's authority would allow and authorize the implementation of arbitrary governance.

14. A Possible Strategy for Fixing the Flaws in the Groundwater Model.

What is really needed, is to correct the flaws in the groundwater model, before the rule is adopted.

One possible strategy for resolving it's problems is be to build a new model, by judiciously drawing from the various earlier studies, while avoiding the more serious mistakes, such as over-parameterization and compromised realism.

The 2003 and 2008 models are unlikely to provide a useful starting point, due to their severe over-parameterization and lack of realism.

In contrast, the 1999 model by Thomas et al. (ibid.) may be more suitable. It comes closer to meeting the standards of science. --- It is relatively realistic and, also, doesn't have many of the problems of the 2003 and 2008 models, because, he directly measured many of its parameters or took them from the literature, rather than fitting them. Nevertheless, there may have been some problems with some of his measurements and he also pointed out, that there remained a few issues that he thought couldn't be resolved. In particular, he was concerned about whether there was subsurface water inflow across the southern boundary of the study area. Nevertheless, as was mentioned above, that, also, remains a concern for the 2003 and 2008 models.

Although, the resulting initial model would probably have relatively low accuracy, it couldn't be any worse than the 2003 and 2008 models, as they have infinite standard errors.

Provided that the resulting model is realistic and has non-infinite confidence intervals, it could provide a starting point for a continuous process of upgrading, using a Bayesian approach. That would allow the incorporation of new data, as it becomes available, as well as the incorporation of new parameter estimates from independent studies, including ones that employed completely different methodologies. For example, in the context of realistic modeling, the parameters have real physical meanings. In that case, it may be possible to directly measure them or to estimate them from independent experiments.

15. Strike out "best available method"

In light of what I have said in comment #14, the phrase "best available method" in WAC 173-518-070 3ai should be removed, as it is undefined and misleading.

16. The 2008 groundwater model might be used for the limited legitimate uses of an empirical model

From a completely different perspective, the 2008 groundwater model could be regarded as an empirical model. In that case, many of the above criticisms are no longer applicable. The legitimate uses of an empirical model are to interpolate short distances within the range of observed data, provided that the underlying processes are known to be consistently applicable throughout that region, but their most appropriate use is to provide an algorithm for the regeneration of data. However, it is being used, as if it were a realistic or theoretical model. In particular, it is being used to make estimates and predictions and to extrapolate beyond the range of the observations or far from them. Those are not valid applications of an empirical model. Nevertheless, there is no reason why it should not be used for any of the *limited* purposes for which it is appropriate. However, that doesn't include most of the types of uses that are involved in supporting the instream flow rule or its implementation.

Nor is there a basis for hope that a continued use of empirical models will lead to an improved understanding.

17. The overall message of the above comments is that the 2008 groundwater flow model is not scientifically valid, it doesn't provide a reasonable basis supporting the rule, it can not reasonably be expected to predict needed mitigation, its outcomes are arbitrary, and it provides the opportunity for arbitrary governance. For these reasons, its use would probably violate due process. Nevertheless, it may be possible to correct its flaws. That needs to be done before the rule is adopted.

Sincerely
Dr. Robert N. Crittenden
June 23, 2012